Theory of Probability:

3rd Sem } Normal laws of error, Standard and Probable error, Least Square fit, Error on the slope and intercept of a fitted line.

5th Hons: Fundamental probability theorem, Conditional probability. Bayes theorem, Repeated trials. Binomial and Multinomial expansion. Random variable and probability distribution. Expectation and variance. Binomial distribution, Poisson Distribution, Continuous distribution: Gaussian or normal distribution.

⊙ Random variables:-

A process is said to be random if it is impossible to predict the final state from the initial state. If a certain group of events is described in a statistical way, it is treated as a random phenomenon. Associated with the random process there may be certain numerical-valued ~~function~~ variables which themselves have random character. ~~Thus~~ These variables are called random variable. They are numerical valued function defined on a sample space.

Repeated Trials:

Sometimes the probability of an event cannot be determined by counting favourable cases and total number of cases. In many such types of cases the probability can be determined by repeating the trials.

Let us consider there are n trials. Each of which is capable of producing one of two possible ~~and~~ outcomes, either a success or a failure.

The success is let denoted by S

& Let the failure is . . . . . F

If the event occur m times,

the relative probability = $\frac{\text{No of occurring the event}}{\text{Total number of trials}}$

$= m/n$

If the event has constant probability P of success, it may be assumed that the

(i) probability of success in each particular trial of the series is the same, denoted by p, and the complementary of success is $q = 1 - p$.

q = the probability of failure of any trial.

(ii) The occurence of success in any particular trial is completely independent of that of all other trials of the series.

If there are m ~~probability~~ of event of success out of n trials, then the number of failure is $(n-m)$

~~To obtain~~ To obtain the number of favourable sequence of m number of success with P probability ~~[and (n−m) number of failure with q probability]~~ in n independent trial, the ~~following~~ required formula is

$$B(m) = {}^{n}C_{m} P^{m} q^{n-m} = \frac{n!}{m!(n-m)!} P^{m} q^{n-m} \longrightarrow (1)$$

which is the $(m+1)^{th}$ term of the binomial expansion of $(P+q)^{n}$

$$(P+q)^{n} = {}^{n}C_{r} P^{r} q^{n-r}$$

If $m = x$

(1) $\rightarrow$ $B(x) = {}^{n}C_{x} P^{x} q^{n-x}$, the Function $B(x)$ gives the binomial frequency function and

And physically $B(x)$ gives the probability of x successes in n independent trials and is called the binomial law of probability.

* As $B(x) = {}^{n}C_{n-x} P^{x} q^{n-x}$ → (ii)

$\therefore B(x-1) = {}^{n}C_{n-(x-1)} P^{x-1} q^{n-(x-1)}$ → (iii)

$\frac{B(x)}{B(x-1)} = \frac{n+1-x}{x} \frac{p}{q} > 1$ for all values of x

$\therefore$ Each probability $B(x)$ is greater than the preceding probability $B(x-1)$ for $x = 1, 2 \ldots n-1$

The Multinomial law:

In binomial law it was considered that the possibility in each of n trails were only two, either any event occur or did not occur. More generally one also meets with the problem that in each of n observations we have more than two possibilities.

Let us consider K mutually exclusive events $A_1, A_2 \cdots A_K$.

Where $A_1 + A_2 \cdots A_K = A$ (certain event) ——— (1)

Let $P_i = P(A_i)$ → the probability that the event A_i occurs.

$$\therefore P_1 + P_2 + \cdots P_K = \sum_{i=1}^{K} P_i = P(A) = 1 \quad \text{——— (2)}$$

Let out of total n $(r_1 + r_2 + \cdots r_K = n)$ trials the events A_1 occures r_1 times, A_2 occurs r_2 times and so on.

If $P_{r_1, r_2 \cdots r_K}$ is the probability among the total ~~trait~~ $r_1 + r_2 + \cdots r_K = n$ trials. These composite event may occur in a number of different ways corresponding to the ν different permutations of $r_1 A_1$'s, $r_2 A_2$'s $\cdots r_K A_K$'s

number

∴ According to compound probability, the ~~probability~~ probability of each one of these permutation is

$$P_1^{r_1} P_2^{r_2} \cdots P_K^{r_K} \quad \text{——— (3)}$$

To get the desired probability $P_{r_1, r_2 \cdots r_k}$, the equ^{n} (3) should be multiplied by the number of ways $r_1, r_2 \cdots r_k$ etc elements can be selected from n ~~events~~ elements.

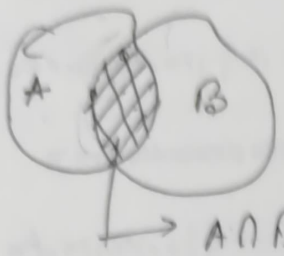
$$\therefore P_{r_1, r_2 \cdots r_k} = \frac{n!}{r_1! r_2! \cdots r_k!} P_1^{r_1} P_2^{r_2} \cdots P_k^{r_k} \rightarrow (4)$$

This formula is called multinomial law.

For $K = 2$, the formula (4) reduced to the binomial law.

Conditional Probability

Let A and B be two subset of a set S in the sample space.



The grilled region according to the set theory can be denoted as A∩B ie A intersection B. The remaining white part in the two domains is A∪B, corresponding to the event set A or B.

The probability of the union A∪B of mutually exclusive events is the sum

$$P(A \cup B) = P(A) + P(B) \quad \text{where } A \cap B = \emptyset$$

For probabilities of arbitrary sets is given by the following relation

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Sometimes the rules and definitions of probabilities discussed above are not sufficient and need to introduce the idea of conditional probability.

Let A & B denote sets of events in one sample space. The conditional probability P(B|A) is defined to be the probability that an event which is member of A is also a member of B.

Mean and Variance :-

We when we make n measurements of a quantity x obtaining the values $x_1, x_2 \cdots x_i \cdots x_n$; the average of the measurements is defined as

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \longrightarrow (1)$$

$\bar{x}$ is called the mean or expectation value, where this formula assumes that every observed value x_i is equally likely and occurs with probability $\frac{1}{n}$.

For discrete random variable x, if x_1 is the measured value with probability p_1, x_2 with corresponding probability p_2 and so on, Then the mean value for the discrete random variable x is

$$\langle X \rangle = \sum_i x_i p_i \qquad \longrightarrow (2)$$

For a continuous random variable x characterised by probability density $f(x)$ as

$$\langle x \rangle = \int x f(x)\, dx \qquad \longrightarrow (3)$$

- There are some other type of mean & such as geometric mean, harmonic mean etc. But the most acceptable one is the arithmetic mean.

Where the sum of quadratic deviations are minimised

~~It gives~~ i.e $\sum_{i=1}^{n}(x-x_i)^2 = 0$ minimum.

$$\Rightarrow 2\left(\sum_i (x-x_i)\right) = 0$$

$$\Rightarrow x = \frac{1}{n}\sum_{i=1}^{n} x_i \equiv \bar{x} \quad \text{---------- (4)}$$

which is nothing but the arithmetic mean.

∴ the arithmetic mean is most accepted one.

Another important property of arithmetic mean

If $v_i = x_i - \bar{x}$, the individual deviations from the mean, Then $\sum_i v_i = 0$

which implies that the sum of positive deviations = Sum of negative deviations.

* The principle of minimizing the quadratic sum of deviation is called the method of least square.

The ability of a mean value to Spread represent a set of data points depends on the spread of the individual measurements from the mean.

A more closer and appropriate definition of the spread is based on the average of the squares of the deviations from the mean. This quantity is known as the standard deviation and mathmetically expressed as.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \longrightarrow (5)$$

Squaring and expanding the above equation we have

$$N\sigma^2 = \sum_{i=1}^{N} x_i^2 - 2\sum_{i} x_i \bar{x} + N\bar{x}^2$$

Letting $\bar{x} \equiv \langle x \rangle$

$$\Rightarrow N\sigma^2 = \sum_{i} x_i^2 - 2\langle x \rangle \sum_{i} x_i + N\langle x \rangle^2$$

$$\Rightarrow N\sigma^2 = N\left(\langle x^2 \rangle - \langle x \rangle^2\right)$$

$$\Rightarrow \sigma^2 = \left(\langle x^2 \rangle - \langle x \rangle^2\right) \text{ where } \sigma^2 \text{ is}$$

correct Called variance.

If all the measured values of x i.e if all x_i are of same value, Then $\langle x^2 \rangle = \langle x \rangle^2$, then variance will be 0, i.e the spread of measurements from the mean value is nil.

$n(S(x-x_i)) = 0$

Random variable

In statistics, Random variable is a type of variable whose possible values depends on the outcomes of a certain random phenomenon. A process is random if it is impossible to predict the final state from the initial state. If a certain group of events is described in a statistical way it is treated as a random phenomenon.

Discrete Random variable:- A discrete Random variable can take only a finite number of distinct values such as 0, 1, 2, 3, ... etc and so on.
Age, height of a person, tossing of coins etc.

~~The values are called~~

Random variables are called random because they can not be defined with certainty. The random phenomena are characterised by the fact that they are conditioned by so many interacting causes that is practically impossible to analyse them accurately to make exact predictions.

The random variables are of two types.
i. Discrete Random variable
& ii. Continuous random variable.

Each outcome of discrete random variable contains a certain probability.

Example: Probability of each outcome (head/tail) of a coin is $1/2$ and the outcomes are of equal probability.

Similarly the probability of each outcome of throwing a dice is $1/6$. The total probability outcome of a discrete variable is equal to 1.

Random variables are denoted by X, Y, Z, etc. A random variable is actually a numerical valued function defined on a sample space.

Symbolically the random variable may be defined as

$$X(e_i) = x_i^\circ, \quad i = 1, 2, 3 \cdots n$$

where e_i are the events of the sample space and x_i° are the values of the random variable X.

(i) If x is a random variable then whatever be the number t, there exists the probability say $P_x(t) = P(x = t)$

(ii) If x is random variable, then whatever be the real no t, there exist the probability $\phi_x(t) = P(-\infty \leq x \leq t)$ that the value assumed by x will not exceed t.

Continuous random variable.

A continuous random variable is one which can take infinite number of possible values. Continuous random variables can be experienced in ~~dandu~~ diffent ——— of daily life.

Some applicatin of these types of ~~rab~~ variables. measuring IQ ~~scor~~ score, Baseball batting average, the length of time in a long distance telephone call, the length of time a computer chip lasts etc.

A Continuous random variable can take on a continuum of values. In other words a random variable is said to be continuous if it assumes a value that falls between a particular interval.

The Central limit Theorem

The central limit theorem states that if a sufficiently large sample is taken from a population, the samples mean will be normally distributed, even if even if the population itself is not normally distributed.

- The mean of the sampling distribution is the mean of the population.

$$\mu_{\bar{x}} = \mu$$

- The standard deviation of the sampling distribution is the standard deviation of the population divided by the square root of the sample size.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Factors that influence the Sampling distribution are the

(i) standard deviation or standard error of the mean and

(ii) variance -

They are mathematical relations and measure

the spread of data points in relation to the mean.

Population

In statistics population is a set of similar items or events which is taken for any statistical study

Samples are a specific collection of group from population. Samples are actually used to make some decision or draw some conclusion about any characteristic of population.

probability sampling means that every member of the target population has a specific chance of being included in the sample.

06/02/25

Fundamental Probability Theorem

*) The following theorems are some fundamental or basic theorem of probability:-

1) The sum of the probability of an event and the probability of not happening the event is always 1.

2) The probability of happening of an impossible event is zero.

3) The probability of an event which will definitely takes place is 1.

4) The probability of any event ~~is~~ lies between 1 and 0.

(5) Bays Theorem related to the The law of total probability describes how to deride a sample space into mutually disjoint set.

6) The addition theorem of probability.

7) The multiplication ~~rule~~ rule of probability.

8) Bays Theorem